Navigation based on Coded Structured Light : Overview and Constraints.

Joaquim Salvi, Joan Batlle and Josep Forest

Computer Vision and Robotics Group
Institute of Informatics and Applications
University of Girona
Avda. Lluis Santalo s/n. 17071 Girona (Spain)
qsalvi@eia.udg.es

Abstract

Structured light has been widely used in mobile robot navigation due to the involved low computing time in segmentation. However, the correspondence problem becomes harder if a complete reconstruction of the scene is required. This paper introduces the coded pattern projection principle which allows to solve completely the matching. Furthermore, focal distances of the imaging and projecting system must vary autonomously with the aim of grabbing always a focused image. This paper overviews the self-calibration concept which arises when a light projector replaces one of both cameras. Finally, the paper proposes the use of a partially self-calibrated system which utilises the advantages of coded pattern projection and camera self-calibration.

1. INTRODUCTION

It is widely known that stereo vision systems are very important to understand the surrounding scene of a mobile robot. In mobile navigation, focal distances of the imaging system must vary autonomously in order to grab always a focused image, and so *a priory* non information of the scene is allowed. This fact leads to the use of an auto-focus system and a method of self-calibration to re-calibrate the system in real-time with no knowledge of the scene. Self-calibration is based on the correspondence matching among the imaging systems. So, most well and precise correspondences are obtained, most precise geometry of the vision system will be computed and most precise 3D information of the scene will be acquired.

The relationship between two cameras may be one of the most used methods. Metric reconstruction of the scene can be obtained from point matching on both image planes. The correspondence of image points is the main problem to be solved in stereo vision. Points without matching or with a false matching lead us to compute a false geometry of the system which will give us a false reconstruction of the scene. This problem is always caused to surface occlusions, projections out of the scope of the camera or simply by a weak algorithm of singular points segmentation.

The problem of image points without matching on the second image plane can be solved by using an stereoscopic system based on structured light. Here, the second stereo camera is replaced by a light source which projects a known pattern on the measuring scene. The first stereo camera images the illuminated scene and analyses the deformations of the imaged pattern with respect to the projected one. Most of the proposed structured light techniques are based on dot or slit patterns. However, the projection of regular patterns involves multiple matching.

In recent years, a new structured light technique has grown in importance. This technique is based on a unique codification for each token of light projected on the scene *[1]*. When the token is imaged by the camera, this codification allows us to obtain the correspondence, that is, to know where it comes from. Then, 2D true matches between camera and projector are directly obtained without using hard computational processes. This technique is widely known as coded structured light or coded pattern projection *[2]*. Although it seems that coded pattern projection could simplify mobile indoor navigation for 3D scene reconstruction *[3]*, there exist several essential problems which do not allow it to be used on self-calibration. This paper will justify the constraints associated with coded pattern projection and will point out that a lot of research is required in this field of computer vision.

The paper is structured as follows. Firstly, the principle of self-calibration and the epipolar geometry is described, explaining how intrinsic and extrinsic parameters can be obtained from stereo vision based on two or most image planes. Secondly, coded pattern projection is introduced focusing in the use of self-calibration. The facility of obtaining the correspondence points from coded structured light becomes an essential problem not allowing the system to be self-calibrated. The paper ends with the conclusions.

2. SELF-CALIBRATION

The technique of self-calibration studies how an imaging system can be calibrated without extracting 3D scene information but only using correspondence matches among image sensors. The relation between the correspondence matches between two cameras can be defined from the epipolar geometry and related by the fundamental matrix *[4, 5]*.

The fundamental matrix relates an image point on an image plane with its epipolar line in the other image plane. Epipolar lines of an image plane coincide in a single point called epipole. As it is shown in figure 1, given an image point m from the first image plane, its correspondence point m' on the second image plane must lie on the epipolar line l'_m i.e. on the plane defined by the three points m, C and C'. The fundamental matrix is then defined from the co-planarity constraint which can be formulated as a product of the vector mC by the vector defined as the cross-product of the vector CC' and the vector $m'C'$.

It has been proved that an image point must lie on the epipolar line defined by its correspondence. Then equation (1) must be satisfied. From the co-planarity constraint the epipolar line can be expressed as (2). Then, the correspondence match relation can be related by a 3x3 matrix F known as the fundamental matrix, as it is expressed in (3) and (4).

$$m' l'_m = 0 \tag{1}$$

$$l'_m = \mathbf{A}^{-T}\mathbf{R}'[t]_x \mathbf{A}'^{-1} m' \tag{2}$$

$$m' \mathbf{F} m' = 0 \tag{3}$$

$$\mathbf{F} = \mathbf{A}^{-T}\mathbf{R}'[t]_x \mathbf{A}'^{-1} \tag{4}$$

In fact, some different fundamental matrices can be computed from a single relationship between two cameras. It depends on where the origin of the world co-ordinate system have been placed and on which image plane the epipolar geometry is computed. In our case, the origin of the co-ordinate system is placed in C and the geometry is computed with respects of I'.

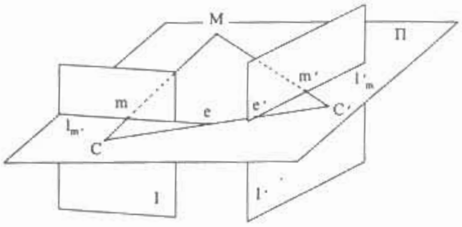


Figure 1. The epipolar geometry which determines the fundamental matrix.

Although the fundamental matrix is a powerful tool in computer vision, more than a single matrix must be computed to infer the intrinsic and extrinsic camera parameters of the imaging system. That is because several possible imaging systems can be modelled from a single **F**. Actually an affine transformation of the whole system is applied. Then three fundamental matrices *[6]* are computed as shown figure 2. $\mathbf{F}_1$ relates the spatial matches of both image planes. $\mathbf{F}_2$ and $\mathbf{F}_3$ relates temporal matches in the same image plane. No restriction in the scene is applied but object movement in order to avoid false matches.

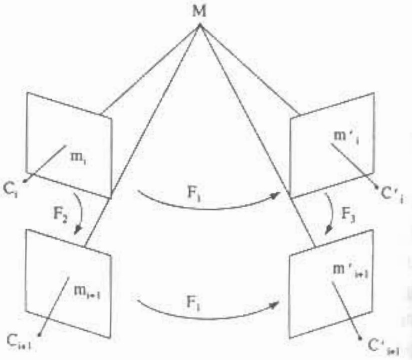


Figure 2. The imaging system movement applied to infer its intrinsic and extrinsic geometrical parameters.

The affine transformation of the imaging system is done by keeping its geometrical relationship then it is supposed that the intrinsic parameters of both imaging systems do not change, neither the rotation matrix **R** and the translation vector t which relates both sensors. In this sense, $\mathbf{F}_2$ and $\mathbf{F}_3$ matrices can be used to compute the intrinsic parameters of both cameras independently. When the intrinsic parameters are known the extrinsic parameters can be computed from $\mathbf{F}_1$, see *[4]* to verify this affirmation.

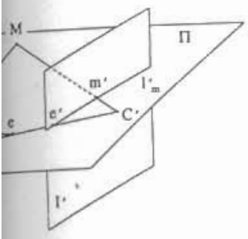
3. CODED PATTERN

One of the most importan
vision systems based on the
cameras is false matching. Fa
concept in stereo vision and h
further computation of the
That is why a stereo vision s
projection of a unique encode
our interesting. In most c
complexity, they patterns
computers and printed by phot
on slides. By this way, each to
by the camera carries th
information to discern where it
matches from surface occlus
matches or points without ma
camera scope are avoided. Alth
control is required, structure
widely used in mobile robot n
pattern projection allows u
correspondence problem witho
time computing algorithms *[1*
systems are based on classical
that is, 3D information of the
from calibrating patterns like r
circles and even using the sam
pattern. This fact restricts the
narrow range of depth distances,
has to be stopped to re-calibrate
distances of the camera and th
must be dynamically readjusted t
best image. We should con
distances do not change in the ca
and the scene measuring step. Of
be changed before it.

It could be thought that the
explained in section 2 could b
calibrate a stereoscopic system
pattern projection. However, an e
should be considered : 3D scen
with no movement to obtain
matches. Then, during the calibra
slide projector must be still, this pr
in figure 3.

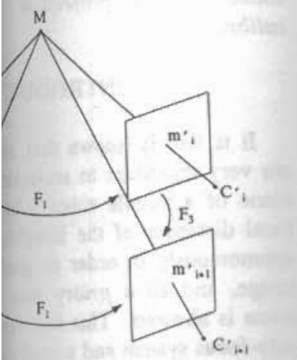
Any movement of the slide proj
calibrating process will produce
the pattern projected on the sc
correspondence is obtained fr
projected, any movement of the
produce a movement of the 3D
carrying an inconsistency on the
fundamental matrix $\mathbf{F}_3$. Furthermo
parameters of the slide projecto
obtained if any movement occurs.
the fundamental matrix $\mathbf{F}_1$ or $\mathbf{F}_2$ ca
but the extrinsic parameters of the
be obtain as we only know the intrin
of the camera sensor meaning

In our case, the origin of
n is placed in C and the
with respects of I'.



geometry which determines the
mental matrix.

lamental matrix is a powerful
on, more than a single matrix
to infer the intrinsic and
meters of the imaging system.
ral possible imaging systems
om a single **F**. Actually an
n of the whole system is
fundamental matrices *[6]* are
igure 2. $\mathbf{F}_1$ relates the spatial
ge planes. $\mathbf{F}_2$ and $\mathbf{F}_3$ relates
the same image plane. No
scene is applied but object
avoid false matches.



ging system movement applied to
extrinsic geometrical parameters.

nsformation of the imaging
by keeping its geometrical
is supposed that the intrinsic
imaging systems do not change,
n matrix **R** and the translation
tes both sensors. In this sense,
s can be used to compute the
eters of both cameras
hen the intrinsic parameters are
ic parameters can be computed
verify this affirmation.

3. CODED PATTERN PROJECTION

One of the most important problems of stereo vision systems based on the relationship of two cameras is false matching. False matching is a key concept in stereo vision and hardly conditions the further computation of the epipolar geometry. That is why a stereo vision system based on the projection of a unique encoded pattern will be of our interesting. In most cases, due to their complexity, they patterns are designed by computers and printed by photographic techniques on slides. By this way, each token of light imaged by the camera carries the required light information to discern where it comes from. False matches from surface occlusions and multiple matches or points without matching due to the camera scope are avoided. Although a scene light control is required, structured light has been widely used in mobile robot navigation. Coded pattern projection allows us to solve the correspondence problem without involving high time computing algorithms *[1]*. However, the systems are based on classical hard calibration, that is, 3D information of the scene is obtained from calibrating patterns like regular squares or circles and even using the same projecting light pattern. This fact restricts the system in a quite narrow range of depth distances, otherwise the task has to be stopped to re-calibrate it. In fact, focal distances of the camera and the slide projector must be dynamically readjusted to grab always the best image. We should consider that focal distances do not change in the calibrating process and the scene measuring step. Of course, they can be changed before it.

It could be thought that the same principle explained in section 2 could be used to self-calibrate a stereoscopic system based on coded pattern projection. However, an essential problem should be considered : 3D scene must be kept with no movement to obtain true temporal matches. Then, during the calibrating process the slide projector must be still, this principle is shown in figure 3.

Any movement of the slide projector during the calibrating process will produce a movement of the pattern projected on the scene. As the correspondence is obtained from the light projected, any movement of the projector will produce a movement of the 3D object points carrying an inconsistency on the compute of the fundamental matrix $\mathbf{F}_3$. Furthermore, the intrinsic parameters of the slide projector can not be obtained if any movement occurs. At this point the fundamental matrix $\mathbf{F}_1$ or $\mathbf{F}_2$ can be computed but the extrinsic parameters of the system can not be obtain as we only know the intrinsic parameters of the camera sensor meaning that a lot of geometrical solutions of the system could be found.

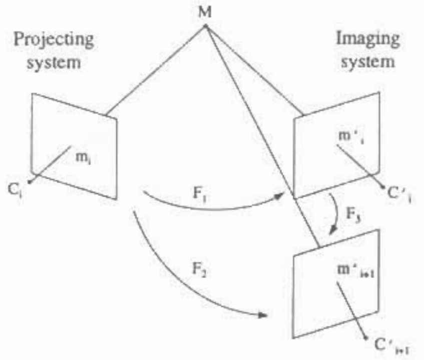


Figure 3. The principle which could be proposed to self-calibrate the pattern projection system.

The problem can be partially solved removing the slide projector with a laser light beam which can be hard calibrated at the beginning and remains always focused. Then the intrinsic parameters of the light projector system will be always known and the camera will be self-calibrated using the principle described in figure 3. However, the 3D scene measurement can not be taken dynamically. During the process the mobile robot should stop. Then, the camera should focused and the $\mathbf{F}_1$ matrix computed. Afterwards, the camera should be moving without changing its internal geometry, that is, its focal, and the $\mathbf{F}_2$ and $\mathbf{F}_3$ matrices computed. Finally, the intrinsic parameters of the camera should be computed and with this a 3D measurement of the scene could be obtained.

Note that if the camera movement is well controlled it is not necessary to compute each time the extrinsic parameters of the system as camera and projecting light system will be always placed at the same positions. This fact, point us to use self-calibration to compute the intrinsic parameters of the camera, the only values which could change.

4. CONCLUSIONS

This paper discussed the problem which arises in self-calibration of stereoscopic systems based on a single imaging sensor and a projector of structured light.

The epipolar geometry of any stereo vision system is obtained from 2D correspondence matches and by the supposition of no movement for 3D object points. This fact causes that stereo vision systems based on structured light can not be self-calibrated. Authors must take into account that a high computational time processing of stereo vision is usually spent to solve the correspondence matching. False matches will always produce geometrical errors which take us to a false epipolar geometry and a consequent rather inaccurate 3D scene reconstruction. Actually, this is the field where coded structured light becomes more important as it can be used to solve the correspondence problem with not a high computational time algorithm, allowing the system to be used in autonomous navigation.

The paper proposes to use a partially self-calibrated system based on the projection of laser light, which utilises the advantages of dynamic camera self-calibration and the advantages of unique matches given by coded pattern projection.

5. REFERENCES

[1] J. Salvi. *An approach to Coded Structured Light to Obtain Three Dimensional Information.* PhD. Thesis University of Girona, December 1997.

[2] J. Batlle, E. Mouaddib and J. Salvi. *Recent progress in Coded Structured Light as a Technique to Solve the Correspondence Problem.* Int. Journal of Pattern Recognition. To be published in 1998.

[3] J. Salvi, E. Mouaddib and J. Batlle. *An Overview of the Advantages and Constraints of Coded Pattern Projection Techniques for Autonomous Navigation.* IEEE Int. Conf. on Intelligent Robots and Systems. Vol. III, pp 1264-1271, 1997.

[4] Q-T. Luong. *Matrice fondamentale et calibration visuelle sur l'environnement vers une plus grande autonomie des systèmes robotiques.* Thèse Université Paris Sud. 1992.

[5] O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint.* The MIT Press 1993, London, England.

[6] Z. Zhang, Q. Loung and O. Faugeras. *Motion of an Uncalibrated Stereo Rig: Self-Calibration and Metric Reconstruction.* Robotics and Automation Journal, Vol. 12, No. 1, pp 103-113, 1996.

Visual rou
localis

Computer Vision and Rob

Abstract

This paper presents a new a
facilitate vision-based autonom
mobile robot on an outdoor
sidewalk. In this sense, we
algorithm for landmark recogniti
The algorithm uses, as landmar
squares where trees have b
recognising module uses techniq
processing, texture analysis and
as well as, some a priori
environment. Three main
presented: Firstly, a robust met
landmarks from the scene is pr
landmarks are located without a
about their allocation, that is, no
navigate. Thirdly, a simple assum
obtain 3D information from a sing

1.Introduction

It is well known that the Sce
subject is an interesting topic in c
be applied to outdoor mobile rob
natural and man-made landmarks
extended technique applied
navigation. Some authors propo
navigation using artificial landm
others, proposed natural landmar
edges [3] or white lines [4,5].

In some cases, a robust extra
landmark features is not an easy t
trees have been used as a robust
for a great number of researchers. F
and Shirai proposed an utiliza
parametric models to interpret natu
on object characteristics, such as
texture [6]. Maeyama, Ohya and Y
method of mobile robot navigatio
landmarks [7]. To detect the